1: Clin Chim Acta 1996 Apr 15;248(1):91-8 Related Articles, Books

Self-organizing neural networks--an alternative way of cluster analysis in clinical chemistry.

Reibnegger G, Wachter H.

Institute of Medical Chemistry, University of Graz, Austria.

Supervised learning schemes have been employed by several workers for training neural networks designed to solve clinical problems. We demonstrate that unsupervised techniques can also produce interesting and meaningful results. Using a data set on the chemical composition of milk from 22 different mammals, we demonstrate that self-organizing feature maps (Kohonen networks) as well as a modified version of error backpropagation technique yield results mimicking conventional cluster analysis. Both techniques are able to project a potentially multi-dimensional input vector onto a two-dimensional space whereby neighborhood relationships remain conserved. Thus, these techniques can be used for reducing dimensionality of complicated data sets and for enhancing comprehensibility of features hidden in the data matrix.

PMID: 8740573 [PubMed - indexed for MEDLINE]

M/S

1: Int J Neural Syst 1999 Jun;9(3):195-202 Related Articles, Books

Estimating the number of clusters in multivariate data by self-organizing maps.

Costa JA, Netto ML.

Department of Computer Engineering and Industry Automation, School of Electrical and Computer Engineering, Universidade Estadual de Campinas, Campinas-SP, Brazil.

Determining the structure of data without prior knowledge of the number of clusters or any information about their composition is a problem of interest in many fields, such as image analysis, astrophysics, biology, etc. Partitioning a set of n patterns in a p-dimensional feature space must be done such that those in a given cluster are more similar to each other than the rest. As there are approximately Kn/K! possible ways of partitioning the patterns among K clusters, finding the best solution is very hard when n is large. The search space is increased when we have no a priori number of partitions. Although the self-organizing feature map (SOM) can be used to visualize clusters, the automation of knowledge discovery by SOM is a difficult task. This paper proposes region-based image processing methods to post-processing the U-matrix obtained after the unsupervised learning performed by SOM. Mathematical morphology is applied to identify regions of neurons that are similar. The number of regions and their labels are automatically found and they are related to the number of clusters in a multivariate data set. New data can be classified by labeling it according to the best match neuron. Simulations using data sets drawn from finite mixtures of p-variate normal densities are presented as well as related advantages and drawbacks of the method.

PMID: 10560758 [PubMed - indexed for MEDLINE]

How many potentially secreted proteins are contained in a bacterial genome?

Schneider G.

F. Hoffmann-La Roche Ltd, Pharmaceuticals Division, Basel, Switzerland. gisbert.schneider@Roche.com

Artificial neural networks were trained on the prediction of the subcellular location of bacterial proteins. A cross-validated average prediction accuracy of 93% was reached for distinction between cytoplasmic and non-cytoplasmic proteins, based on the analysis of protein amino-acid composition. Principal component analysis and self-organizing maps were used to create graphical representations of amino-acid sequence space. A clear separation of cytoplasmic, periplasmic, and extracellular proteins was observed. The neural network system was applied to predicting potentially secreted proteins in 15 complete genomes. For mesophile bacteria the predicted fractions of non-cytoplasmic proteins agree with previously published estimates, ranging between 15% and 30%. Characteristics of thermophile genomes might lead to an under-estimation of the fraction of secreted proteins by presently available prediction systems. A self-organizing map was constructed from all 15 bacterial genomes. This technique can reveal additional sequence features independent from exhaustive pair-wise sequence alignment. The Treponema pallidum and Mycobacterium tuberculosis data formed separate clusters indicating unusual characteristics of these genomes.

PMID: 10524242 [PubMed - indexed for MEDLINE]

1: Bioinformatics 1999 Sep;15(9):741-8 Related Articles, Books, LinkOut

Associative database of protein sequences.

Hanke J, Lehmann G, Bork P, Reich JG.

Max-Delbruck-Center for Molecular Medicine, Department of Bioinformatics, Robert-Rossle-Strasse 10, D-13125 Berlin-Buch, Germany.

MOTIVATION: We present a new concept that combines data storage and data analysis in genome research, based on an associative network memory. As an illustration, 115 000 conserved regions from over 73 000 published sequences (i.e. from the entire annotated part of the SWISSPROT sequence database) were identified and clustered by a self-organizing network. Similarity and kinship, as well as degree of distance between the conserved protein segments, are visualized as neighborhood relationship on a two-dimensional topographical map. RESULTS: Such a display overcomes the restrictions of linear list processing and allows local and global sequence relationships to be studied visually. Families are memorized as prototype vectors of conserved regions. On a massive parallel machine, clustering and updating of the database take only a few seconds; a rapid analysis of incoming data such as protein sequences or ESTs is carried out on present-day workstations. AVAILABILITY: Access to the database is available at http://www.bioinf.mdc-berlin.de/unter2.html++ + CONTACT: (hanke,lehmann,reich)@mdc-berlin.de; bork@embl-heidelberg.de

PMID: 10498774 [PubMed - indexed for MEDLINE]

1: Eur J Clin Chem Clin Biochem 1993 May;31(5):311-6

Related Articles, Books

Self-organizing neural networks as a means of cluster analysis in clinical chemistry.

Reibnegger G, Weiss G, Wachter H.

Institute of Medical Chemistry and Biochemistry, University of Innsbruck, Austria.

Connectionist systems (often termed "neural networks") are an alternative way to solve data processing tasks. They differ radically from conventional "von-Neumann" computing devices. Recent work on neural networks in clinical chemistry was done using supervised learning schemes, resulting in models which resemble classical discriminant analysis. The aim of the present study is to make clinical chemists familiar with basic concepts of self-organizing neural networks employing unsupervised learning schemes. Using a benchmark data set on the composition of milk from 22 different mammals, it is demonstrated that self-organizing neural networks are capable of performing tasks similar to classical cluster analysis and principal component analysis. Self-organizing neural networks could be envisaged to provide an alternative way for reducing the dimensionality of complex multivariate data sets, thus producing easily comprehensible low-dimensional "maps" of essential features.

PMID: 8357940 [PubMed - indexed for MEDLINE]

1: Subst Use Misuse 1998 Jan;33(2):365-81 Related Articles, Books

Self-organizing maps.

Matera F.

Semeion Research Center, Rome, Italy.

PMID: 9516733 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

A neural network approach to automatic chromosome classification.

Jennings AM, Graham J.

Department of Medical Biophysics, University of Manchester, UK.

Classification of banded metaphase chromosomes is an important step in automated clinical chromosome analysis. We have conducted a preliminary investigation of the application of artificial neural networks to this process, making use of a natural representation of the banding pattern. Two different network architectures have been compared: the Kohonen self-organizing feature map and the multi-layer perception (MLP). For each of these a search of their respective parameter spaces over a limited range has resulted in configurations of modest dimension which achieve creditable classification rates. The MLP in particular shows promise of being a useful classifier. When size and shape features are supplied as inputs to the MLP in addition to a low-resolution banding profile, misclassification rates are obtained which are comparable with those of a well developed statistical classifier.

PMID: 8372108 [PubMed - indexed for MEDLINE]

1: Biophys J 1994 Jun;66(6):1804-14 Related Articles, Books

Pattern recognition and classification of images of biological macromolecules using artificial neural networks.

Marabini R, Carazo JM.

Centro Nacional de Biotecnologia (CSIC), Universidad Autonoma, Madrid, Spain.

The goal of this work was to analyze an image data set and to detect the structural variability within this set. Two algorithms for pattern recognition based on neural networks are presented, one that performs an unsupervised classification (the self-organizing map) and the other a supervised classification (the learning vector quantization). The approach has a direct impact in current strategies for structural determination from electron microscopic images of biological macromolecules. In this work we performed a classification of both aligned but heterogeneous image data sets as well as basically homogeneous but otherwise rotationally misaligned image populations, in the latter case completely avoiding the typical reference dependency of correlation-based alignment methods. A number of examples on chaperonins are presented. The approach is computationally fast and robust with respect to noise. Programs are available through ftp.

PMID: 7915552 [PubMed - indexed for MEDLINE]

1: Ultramicroscopy 1996 Sep;65(1-2):81-93 Related Articles, Books

Analysis of structural variability within two-dimensional biological crystals by a combination of patch averaging techniques and self organizing maps.

Fernandez JJ, Carazo JM.

Centro Nacional de Biotecnologia-CSIC, Universidad Autonoma, Madrid, Spain.

We study in this work the use of self organizing maps to analyze the structural variability that can be found along two-dimensional crystals of biological macromolecules. Small areas of the crystals, termed "patches" by previous researchers, are used to obtain local average images that are then used as the input of a Self Organizing Map. This procedure allows for a fast and accurate image classification. Multivariate Statistical Analysis is then used on the resulting code vectors producing a very condensed data representation. This methodology is applied to previously studied crystals of bacteriophage phi 29 p10 connector, finding a crystalline heterogeneity probably associated to multilayers in some areas of the crystal.

PMID: 8961548 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

1: Electrophoresis 1996 Apr;17(4):694-8 Related Articles, Books

From image processing to classification: IV. Classification of electrophoretic patterns by neural networks and statistical methods enable quality assessment of wheat varieties for breadmaking.

Jensen K, Kesmir C, Sondergaard I.

Department of Biochemistry and Nutrition, Technical University of Denmark, Lyngby, Denmark.

The end-use quality of products made from doughs consisting of wheat flour and water is often dependent upon the storage (gluten) proteins of the grain endosperm. Today the electrophoretic patterns of the high molecular weight (HMW) glutenin subunits are used for quality selections in wheat breeding programs in several countries. In this study, we used two multivariate techniques to classify digitized patterns from isoelectric focusing of gliadins and glutenins: a two-layered neural network architecture consisting of a self-organizing feature map and a feed-forward classifier [1], and discriminant analysis [2,3]. Three groups of seven wheat varieties (Triticum aestivum L.), associated with poor, medium or good properties in relation to bread-making quality, were used. The best classification results were obtained by the neural network model, based on data from the gliadin fraction: it was possible to classify varieties associated with poor or good quality, with recognition rates of 70 and 69%, respectively. The statistical method was better suited to solve the classification problem when the data was based on the glutenin fraction: if a specific variety was already known to be non-poor, this method enabled us to classify the medium- and good-quality classes with recognition rates of 90 and 88%, respectively. The results obtained were confirmed by correlation coefficients.

PMID: 8738329 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

1: Neural Comput 1995 Nov;7(6):1188-90 Related Articles, Books

Sorting with self-organizing maps.

Budinich M.

INFN, Trieste, Italy.

A self-organizing feature map (Von der Malsburg 1973; Kohonen 1984) sorts n real numbers in O(n) time apparently violating the O(n log n) bound. Detailed analysis shows that the net takes advantage of the uniform distribution of the numbers and, in this case, sorting in O(n) is possible. There are, however, an exponentially small fraction of pathological distributions producing O(n2) sorting time. It is interesting to observe that standard learning produced a smart sorting algorithm.

PMID: 7584897 [PubMed - indexed for MEDLINE]

1: Biol Cybern 1997 Jun;76(6):441-50 Related Articles, Books, LinkOut

Classification of protein families and detection of the determinant residues with an improved self-organizing map.

Andrade MA, Casari G, Sander C, Valencia A.

Protein Design Group, Centro Nacional de Biotecnologia-CSIC, Cantoblanco, Madrid, Spain. andrade@ebi.ac.uk

Using a SOM (self-organizing map) we can classify sequences within a protein family into subgroups that generally correspond to biological subcategories. These maps tend to show sequence similarity as proximity in the map. Combining maps generated at different levels of resolution, the structure of relations in protein families can be captured that could not otherwise be represented in a single map. The underlying representation of maps enables us to retrieve characteristic sequence patterns for individual subgroups of sequences. Such patterns tend to correspond to functionally important regions. We present a modified SOM algorithm that includes a convergence test that dynamically controls the learning parameters to adapt them to the learning set instead of being fixed and externally optimized by trial and error. Given the variability of protein family size and distribution, the addition of this features is necessary. The method is successfully tested with a number of families. The rab family of small GTPases is used to illustrate the performance of the method.

PMID: 9263431 [PubMed - indexed for MEDLINE]

1: J Natl Cancer Inst 1994 Dec 21;86(24):1853-9 Related Articles, Books

Use of the Kohonen self-organizing map to study the mechanisms of action of chemotherapeutic agents.

van Osdol WW, Myers TG, Paull KD, Kohn KW, Weinstein JN.

Laboratory for Molecular Pharmacology, National Cancer Institute, Bethesda, Md 20892.

BACKGROUND: Many natural and synthetic compounds might prove to be effective in cancer chemotherapy. To identify potentially useful agents, the National Cancer Institute screens over 10,000 compounds annually against a panel of 60 distinct human tumor cell lines in vitro. This screening program generates large amounts of data that are organized into relational databases. Important questions concern the information content of the data and ways to extract that information. Previously, statistical techniques have revealed that compounds with similar patterns of activity against the 60 cell lines are often similar in structure and mechanism of action. Feed-forward, back-propagation neural networks have been trained on this type of data to predict broadly defined mechanisms of action of chemotherapeutic agents. PURPOSE AND METHOD: In this report, we examine the information that can be extracted from the screening data by means of another type of neural network paradigm, the Kohonen self-organizing map. This is a topology-preserving function, obtained by unsupervised learning, that nonlinearly projects the high-dimensional activity patterns into two dimensions. Our dataset is almost identical to that used in the earlier neural network study. RESULTS: The self-organizing maps we constructed have several important characteristics. 1) They partition the two-dimensional array into distinct regions, each of which is principally occupied by agents having the same broadly defined mechanism of action. 2) These regions can be resolved into distinct subregions that conform to plausible submechanisms and chemically defined subgroups of submechanism. 3) These results (and exceptions to them) are consistent with those obtained with the use of such deterministic measures of similarity among activity patterns as the Euclidean distance or Pearson correlation coefficient. CONCLUSIONS: Our results indicate that the activity patterns obtained from the screen contain detailed information about mechanism of action and its basis in chemical structure. The self-organizing map can be used

to suggest the mechanism of action of compounds identified by the screen as potentially useful chemotherapeutic agents and to probe the biology of the cell lines in the cancer screen. Kohonen self-organizing maps, unlike the previously applied neural networks, preserve and reveal the relationships among compounds acting by similar mechanisms and therefore have the potential to identify compounds that act by novel cytotoxic mechanisms.

PMID: 7990160 [PubMed - indexed for MEDLINE]

1: Neural Comput 1995 Nov;7(6):1188-90 Related Articles, Books

Sorting with self-organizing maps.

Budinich M.

INFN, Trieste, Italy.

A self-organizing feature map (Von der Malsburg 1973; Kohonen 1984) sorts n real numbers in O(n) time apparently violating the O(n log n) bound. Detailed analysis shows that the net takes advantage of the uniform distribution of the numbers and, in this case, sorting in O(n) is possible. There are, however, an exponentially small fraction of pathological distributions producing O(n2) sorting time. It is interesting to observe that standard learning produced a smart sorting algorithm.

PMID: 7584897 [PubMed - indexed for MEDLINE]

1: Biosystems 1997;41(2):105-25 Related Articles, Books

Learning systems in biosignal analysis.

Schizas CN, Pattichis CS.

Department of Computer Science, University of Cyprus, Nicosia.
schizas@turing.cs.ucy.ac.cy

In biosignal analysis, the utility of artificial neural networks (ANN) in classifying electromyographic (EMG) data trained with the momentum back propagation algorithm has recently been demonstrated. In the current study, the self-organizing feature map algorithm, the genetics-based machine learning (GBML) paradigm, and the K-means nearest neighbour clustering algorithm are applied on the same set of data. The aim of this exercise is to show how these three paradigms can be used in practice, given that their diagnostic performance is problem- and parameter-dependent. A total of 720 macro EMG recordings were carried out from four groups, from seven normal, nine motor neuron disease, 14 Becker's muscular dystrophy, and six spinal muscular atrophy subjects, respectively. Twenty-three of the subjects were used for training and 13 for evaluating the various models. For each subject, the mean and the standard deviation of the parameters (i) amplitude, (ii) area, (iii) average power and (iv) duration were extracted. The feature vector was structured in two different ways for input to the models: an eight-input feature vector that consisted of both the mean and the standard deviation of the four parameters measured, and a four-input feature vector that included only the mean of the parameters. Also, due to the heterogenous nature of the spinal muscular atrophy group, three class models that excluded this group were investigated. In general, self-organizing feature map and GBML models resulted in comparable diagnostic performance of the order of 80-90% correct classifications (CCs) score for the evaluation set, whereas the K-means nearest neighbour algorithm models gave lower percentage CCs. Furthermore, for all three learning paradigms: better diagnostic performance was obtained for the three class models compared with the four class models; similar diagnostic performance was obtained for both the eight- and four-input feature vectors. Finally, it is claimed that the proposed methodology followed in this work can be applied for the development of diagnostic systems in the analysis of biosignals.

M/S

PMID: 9043680 [PubMed - indexed for MEDLINE]

Full Record

Return to PNAS

A genome-wide transcriptional analysis of the mitotic cell cycle

Cho, RJ;Campbell, MJ;Winzeler, EA;Steinmetz, L;Conway, A;Wodicka, L;Wolfsberg, TG;Gabrielian, AE;Landsman, D;Lockhart, DJ;Davis, RW

MOLECULAR CELL
2: (1) 65-73 JUL 1998

Document type: Article **Language:** English

Abstract:
Progression through the eukaryotic cell cycle is known to be both regulated and accompanied by periodic fluctuation in the expression levels of numerous genes. We report here the genome-wide characterization of mRNA transcript levels during the cell cycle of the budding yeast S. cerevisiae. Cell cycle-dependent periodicity was found for 416 of the 6220 monitored transcripts. More than 25% of the 416 genes were found directly adjacent to other genes in the genome that displayed induction in the same cell cycle phase, suggesting a mechanism for local chromosomal organization in global mRNA regulation. More than 60% of the characterized genes that displayed mRNA fluctuation have already been implicated in cell cycle period-specific biological roles. Because more than 20% of human proteins display significant homology to yeast proteins, these results also link a range of human genes to cell cycle period-specific biological functions.

Addresses:
Campbell MJ, Mol Applicat, 607 Hansen Way, Bldg 1, Palo Alto, CA 94304 USA.
Stanford Univ, Sch Med, Dept Genet, Stanford, CA 94305 USA.
Stanford Univ, Sch Med, Dept Biochem, Stanford, CA 94305 USA.
Affymetrix, Santa Clara, CA 95051 USA.
Natl Lib Med, Natl Ctr Biotechnol Informat, Bethesda, MD 20894 USA.

Authors' E-mail Addresses:

Publisher:
CELL PRESS, CAMBRIDGE

TGA Number:
106WR

ISSN:
1097-2765

Return to PNAS

1: Proc Natl Acad Sci U S A 1999 Mar 16;96(6):2907-12

Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation.

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR.

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA.

Array technologies have made it straightforward to monitor simultaneously the expression pattern of thousands of genes. The challenge now is to interpret such massive data sets. The first step is to extract the fundamental patterns of gene expression inherent in the data. This paper describes the application of self-organizing maps, a type of mathematical cluster analysis that is particularly well suited for recognizing and classifying features in complex, multidimensional data. The method has been implemented in a publicly available computer package, GENECLUSTER, that performs the analytical calculations and provides easy data visualization. To illustrate the value of such analysis, the approach is applied to hematopoietic differentiation in four well studied models (HL-60, U937, Jurkat, and NB4 cells). Expression patterns of some 6,000 human genes were assayed, and an online database was created. GENECLUSTER was used to organize the genes into biologically relevant clusters that suggest novel hypotheses about hematopoietic differentiation-for example, highlighting certain genes and pathways involved in "differentiation therapy" used in the treatment of acute promyelocytic leukemia.

PMID: 10077610 [PubMed - indexed for MEDLINE]

1: FEBS Lett 1999 May 21;451(2):142-6 Related Articles, Books

Analysis of gene expression data using self-organizing maps.

Toronen P, Kolehmainen M, Wong G, Castren E.

A.I. Virtanen Institute, University of Kuopio, Finland.

DNA microarray technologies together with rapidly increasing genomic sequence information is leading to an explosion in available gene expression data. Currently there is a great need for efficient methods to analyze and visualize these massive data sets. A self-organizing map (SOM) is an unsupervised neural network learning algorithm which has been successfully used for the analysis and organization of large data files. We have here applied the SOM algorithm to analyze published data of yeast gene expression and show that SOM is an excellent tool for the analysis and visualization of gene expression profiles.

PMID: 10371154 [PubMed - indexed for MEDLINE]

Display Abstract Save Text Order Add to Clipboard

DERWENT-ACC-NO: 2000-573798
DERWENT-WEEK: 200104

TITLE: Clustering gene expression datapoints in a computer system using a
self-organizing map
INVENTOR: GOLUB, T R; LANDER, E S ; MESIROV, J ; TAMAYO, P
PRIORITY-DATA: 1999US-0124453 (March 15, 1999)
PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
JP 2000342299 A	December 12, 2000	N/A	034	C12Q 001/68
	September 20, 2000	E	039	G06F 019/00
EP 1037158 A2	September 15, 2000	E	000	G06F 019/00
CA 2300639 A1				

INT-CL (IPC): C12M001/00; C12N001/00 ; C12N015/09 ; C12Q001/68 ;
G01N033/15 ; G01N033/50 ; G06F017/30 ; G06F019/00
ABSTRACTED-PUB-NO: EP 1037158A
BASIC-ABSTRACT: NOVELTY - A method for clustering gene expression datapoints in
a computer system using a self-organizing map, is new.

DETAILED DESCRIPTION - Method for clustering datapoints (each datapoint is a
series of gene expression values) in a computer system, comprises:

(a) receiving the gene expression values of the datapoints;

(b) using a self-organizing map (SOM), clustering the datapoints so that
datapoints that exhibit similar patterns are clustered together into respective
clusters; and

(c) providing an output indicating the clusters of the datapoints.

INDEPENDENT CLAIMS are also included for the following:

(1) a method for grouping datapoints in a computer system, where each datapoint
is a series of gene expression values, comprising:

(i) receiving gene expression values of the datapoints;

(ii) filtering out any datapoints that exhibit an insignificant change in the
gene expression value, so that working datapoints remain;

(iii) normalizing the gene expression value of the working datapoints;

(iv) using a SOM, grouping the working datapoints so that datapoints that
exhibit similar patterns are grouped together into respective clusters; and

(v) providing an output indicating the groups of the datapoints;

(2) a computer apparatus for clustering datapoints, where each datapoint is a
series of gene expression values, comprising:

(i) a source of gene expression values of the datapoints;

(ii) a processor routine coupled to receive datapoints from the source, the
processor routine utilizing a SOM for clustering datapoints so that datapoints
that exhibit similar patterns are clustered together into respective clusters;
and

(iii) an output device, coupled to the processor routine, for indicating the
clusters of datapoints;

(3) a computer apparatus for grouping datapoints, where each datapoint is a
series of gene expression values, comprising:

(i) a source of gene expression values of the datapoints;

(ii) a filter coupled to the source, for receiving the gene expression values
and filtering out any of the datapoints that exhibit an insignificant change in
the gene expression value, so that working datapoints remain;

(iii) a normalizing process, coupled to the filter, for normalizing the gene
expression value of the working datapoints;

(iv) a processor routine that is responsive to the normalizing process and
utilizes a SOM for grouping the working datapoints such that datapoints that
exhibit similar patterns are grouped together into respective groups; and

(v) an output device, coupled to the processor routine, for indicating the
clusters of datapoints;

(4) a method for assessing expression patterns of two or more genes in

cells, where the expression patterns are represented by datapoints, and each
datapoint is a series of gene expression values, comprising:

(i) receiving the gene expression values of the datapoints;

(ii) using a SOM, clustering the datapoints such that datapoints that exhibit
similar patterns are clustered together into respective clusters;

(iii) providing an output indicating the clusters of datapoints; and

(iv) analyzing the output to determine the similarities or differences between
the expression patterns of the genes;

(5) a method of determining relatedness of expression patterns of two or more
genes, where the expression patterns are represented by datapoints and each
datapoint is a series of gene expression values, comprising:

(i) receiving the gene expression values of the datapoints;

(ii) using a SOM, clustering the datapoints such that datapoints that exhibit
similar patterns are clustered together into respective clusters;

(iii) providing an output indicating the clusters of datapoints; and

(iv) analyzing the output to determine the similarities and/or differences
between the expression patterns of the genes, thereby determining the

relatedness of the genes;

(6) a method for characterizing expression patterns of genes of a sample having unknown characteristics, where the sample is obtained from an individual and subjected to diagnostic tests, and the expression patterns of the genes for the diagnostic tests are represented by datapoints, and each datapoint is a series of gene expression values across multiple genes for the diagnostic test, comprising:

(i) receiving the gene expression values of the datapoints from the diagnostic tests;

(ii) using a SOM, clustering the datapoints such that datapoints that exhibit similar patterns are clustered together into respective clusters;

(iii) providing an output indicating the clusters of datapoints; and

(iv) comparing the output of the gene expression patterns of the unknown sample against a control, thereby characterizing gene expression patterns of the sample;

(7) a method of identifying a drug target from the expression patterns of two or more genes from cells, where the expression patterns are represented by datapoints and each datapoint is a series of gene expression values, comprising:

(i) obtaining cells that express genes;

(ii) subjecting the cells to an agent or condition for testing the drug target;

(iii) measuring gene expression from the cells subjected to the agent or condition, and from a control, to obtain the gene expression values;

(iv) receiving the gene expression values of the datapoints;

(v) using a SOM, clustering the datapoints such that datapoints that exhibit
similar patterns are clustered together into respective clusters;

(vi) comparing the clusters from the genes that have been subjected to the
agents or condition with a control; and

(vii) providing an output indicating clusters, to thereby determine the drug
target;

(8) a drug target identified or identifiable by the method of (7);

(9) a computer-readable product on which is recorded a program loadable into
the internal memory of a digital computer and comprising software code portions
for performing the steps of the above methods.

USE - The method can be used, e.g. to identify drug targets from the expression
patterns of two or more genes and to analyze the relatedness of two or more
genes, the unknown function of a gene under known conditions, the effect of
unknown conditions on a known gene function or the likelihood of successful
treatment by a drug (e.g. for a specific tissue sample).

ADVANTAGE - Using SOMs to cluster gene expression patterns into groups
exhibiting similar patterns makes it easy to analyze gene expression data from
potentially thousands of genes.

DESCRIPTION OF DRAWING(S) - The figure is a schematic diagram illustrating the
principle behind the self-organizing map, in which the initial geometry of
nodes in a 3x2 rectangular grid is indicated by solid lines connecting the
nodes, datapoints are represented by black dots, the nodes are represented by
large circles, and trajectories are represented by arrows.
CHOSEN-DRAWING: Dwg.1/6

L5 ANSWER 1 OF 2 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1994:104157 BIOSIS
DOCUMENT NUMBER: PREV199497117157
TITLE: Potentially functional regions of **nucleic** acids recognized by a Kohonen's **self-organizing map**.
AUTHOR(S): Giuliano, F.; Arrigo, P.; Scalia, F.; Cardo, P. P.; Damiani, G. (1)
CORPORATE SOURCE: (1) Istituto Policattedra di Chimica Biologica, Viale Benedetto XV 1, 16232 Genova Italy
SOURCE: Computer Applications in the Biosciences, (1993) Vol. 9, No. 6, pp. 687-693.
ISSN: 0266-7061.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:
Computer recognition of short functional sites on DNA, such as promoter regions or intron-exon boundaries, has recently attracted much interest. In this paper we have focused our attention on the automatic recognition of relevant features of human **nucleic** acid sequences by means of an unsupervised artificial neural network model. Sixty messenger RNA and 31 genomic DNA sequences were analysed. The results showed that in mRNA, the minimal similarity 60 base pattern was guanine- and cytosine-rich and located in most sequences in a range of 250 bases from either the middle point of the signal peptide coding region or from the start of the coding region. On DNA sequences a region defined by a cluster of minimal similarity patterns was present in many of the analysed genes. This zone may be related to alternative splicing and DNA methylation.
CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Nucleic Acids, Purines and Pyrimidines *10052
Biochemical Studies - Nucleic Acids, Purines and Pyrimidines 10062
Biophysics - Molecular Properties and Macromolecules *10506
Nervous System - General; Methods *20501
INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Information Studies; Mathematical Biology (Computational Biology); Methods and Techniques; Nervous System (Neural Coordination)
INDEX TERMS: Sequence Data
nucleotide sequence
INDEX TERMS: Miscellaneous Descriptors
ARTIFICIAL NEURAL NETWORK; COMPUTER ANALYSIS; DNA; EMBL; METHYLATION; SPLICING

L5 ANSWER 2 OF 2 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1991:431331 BIOSIS
DOCUMENT NUMBER: BA92:87496
TITLE: IDENTIFICATION OF A NEW MOTIF ON **NUCLEIC** ACID SEQUENCE DATA USING KOHONEN'S **SELF-ORGANIZING MAP**.
AUTHOR(S): ARRIGO P; GIULIANO F; SCALIA F; RAPALLO A; DAMIANI G
CORPORATE SOURCE: IST. I CIRCUITI ELETTRONICI C.N.R., VIA ALL'OPERA PIA 11, 16145 GENOVA, ITALY.
SOURCE: COMPUT APPL BIOSCI, (1991) 7 (3), 353-358.
CODEN: COABER. ISSN: 0266-7061.
FILE SEGMENT: BA; OLD
LANGUAGE: English
ABSTRACT:
Here we present a performance test of a Kohonen features **map** applied to the fast extraction of uncommon sequences from the coding region of the

human insulin receptor gene. We used a network with 30 neurons and with a variable input window. The program was aimed at detecting unique or uncommon DNA regions present in crude sequence data and was able to automatically detect the signal peptide coding regions of a set of human insulin receptor gene data. The testing of this programn with HSIRPR cDNA release (EMBL data bank) indicated the presence of unique features in the signal peptide coding region. On the basis of our results this program can automatically detect 'singularity' from crude sequencing data and it does not require knowledge of the features to be found.

CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Methods, Materials and Apparatus, General - Laboratory Apparatus 01006
Genetics and Cytogenetics - Human *03508
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Nucleic Acids, Purines and Pyrimidines 10052
Biophysics - Bioengineering 10511
Biophysics - Biocybernetics *10515
Metabolism - Nucleic Acids, Purines and Pyrimidines *13014
Endocrine System - Pancreas *17008
Psychiatry - Mental Retardation 21006

BIOSYSTEMATIC CODE: Hominidae 86215

INDEX TERMS: Miscellaneous Descriptors
HUMAN INSULIN RECEPTOR GENE DATA

REGISTRY NUMBER: 9004-10-8 (INSULIN)

L6 ANSWER 1 OF 30 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 2000:407420 BIOSIS
DOCUMENT NUMBER: PREV200000407420
TITLE: Hierarchical state space partitioning with a network self-organising **map** for the recognition of ST-T segment changes.
AUTHOR(S): Bezerianos, A. (1); Vladutu, L.; Papadimitriou, S.
CORPORATE SOURCE: (1) Department of Medical Physics, School of Medicine, University of Patras, Patras Greece
SOURCE: Medical & Biological Engineering & Computing, (July, 2000) Vol. 38, No. 4, pp. 406-415. print.
ISSN: 0140-0118.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:
The problem of maximising the performance of ST-T segment automatic recognition for ischaemia detection is a difficult **pattern** classification problem. The paper proposes the network self-organising **map** (NetSOM) model as an enhancement to the Kohonen self-organised **map** (SOM) model. This model is capable of effectively decomposing complex large-scale ***pattern*** classification problems into a number of partitions, each of which is more manageable with a local classification device. The NetSOM attempts to generalise the regularisation and ordering potential of the basic SOM from the space of vectors to the space of approximating functions. It becomes a device for the ordering of local experts (i.e. independent neural networks) over its lattice of neurons and for their selection and co-ordination. Each local expert is an independent neural network that is trained and activated under the control of the NetSOM. This method is evaluated with examples from the European ST-T database. The first results obtained after the application of NetSOM to ST-T segment change recognition show a significant improvement in the performance compared with that obtained with monolithic approaches, i.e. with single network types. The basic SOM model has attained an average ischaemic beat sensitivity of 73.6% and an average ischaemic beat predictivity of 68.3%. The work reports and discusses the improvements that have been obtained from the implementation of a NetSOM classification system with both multilayer perceptrons and radial basis function (RBF) networks as local experts for the ST-T segment change problem. Specifically, the NetSOM with multilayer perceptrons (radial basis functions) as local experts has improved the results over the basic SOM to an average ischaemic beat sensitivity of 75.9% (77.7%) and an average ischaemic beat predictivity of 72.5% (74.1%).
CONCEPT CODE: Nervous System - Physiology and Biochemistry *20504
Cytology and Cytochemistry - Animal *02506
Biophysics - Bioengineering *10511
INDEX TERMS: Major Concepts
Biomedical Engineering (Allied Medical Sciences); Nervous System (Neural Coordination)
INDEX TERMS: Parts, Structures, & Systems of Organisms
neural network: nervous system; neuron: lattice, nervous system
INDEX TERMS: Methods & Equipment
hierarchical state space partitioning: analytical method
INDEX TERMS: Miscellaneous Descriptors
Kohonen **self-organizing map**
model; ST-T segment changes: automatic recognition; ischemic beat predictivity; multilayer perceptrons; network **self-organizing map**

L6 ANSWER 6 OF 30 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1999:176708 BIOSIS
DOCUMENT NUMBER: PREV199900176708
TITLE: A neural network approach to the analysis and classification of human craniofacial growth.
AUTHOR(S): Lux, C. J. (1); Stellzig, A.; Volz, D.; Jaeger, W.; Richardson, A.; Komposch, G.
CORPORATE SOURCE: (1) Department of Orthodontics, Dental School, University of Heidelberg, Im Neuenheimer Feld 400, 69120, Heidelberg Germany
SOURCE: Growth Development and Aging, (Autumn, 1998) Vol. 62, No. 3, pp. 95-106.
ISSN: 1041-1232.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:
Planning of treatment in the field of orthodontics and maxillo-facial surgery is largely dependent on the individual growth of a patient. In the present work, the growth of 43 orthodontically untreated children was analysed by means of lateral cephalograms taken at the ages of 7 and 15. For the description of craniofacial skeletal changes, the concept of tensor analysis and related methods have been applied. Thus the geometric and analytical shortcomings of conventional cephalometric methods have been avoided. Through the use of an artificial neural network, namely **self-organizing** neural maps, the resultant growth data were classified and the relationships of the various growth patterns were monitored. As a result of self-organization, the 43 children were topologically ordered on the emerging **map** according to their craniofacial size and shape changes during growth. As a new patient can be allocated on the **map**, this type of network provides a frame of reference for classifying and analysing previously unknown cases with respect to their growth **pattern**. If landmarks are used for the determination of growth, the morphometric methods applied as well as the subsequent visualization of the growth data by means of neural networks can be employed for the analysis and classification of growth-related skeletal changes in general.
CONCEPT CODE: Bones, Joints, Fasciae, Connective and Adipose Tissue - Physiology and Biochemistry *18004
General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Chordate Body Regions - Head *11304
Chordate Body Regions - Facial *11306
Developmental Biology - Embryology - Morphogenesis, General *25508
Mathematical Biology and Statistical Methods *04500
Nervous System - General; Methods *20501
BIOSYSTEMATIC CODE: Hominidae 86215
INDEX TERMS: Major Concepts
Computer Applications (Computational Biology); Development; Orthopedics (Human Medicine, Medical Sciences)
INDEX TERMS: Miscellaneous Descriptors
craniofacial growth: analysis, classification, computer program, neural network approach
ORGANISM: Super Taxa
Hominidae: Primates, Mammalia, Vertebrata, Chordata, Animalia
ORGANISM: Organism Name
human (Hominidae): normal subjects
ORGANISM: Organism Superterms
Animals; Chordates; Humans; Mammals; Primates; Vertebrates

L6 ANSWER 8 OF 30 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1998:482703 BIOSIS
DOCUMENT NUMBER: PREV199800482703
TITLE: Profiles of chemically-induced tumors in rodents: Quantitative relationships.
AUTHOR(S): Benigni, Romualdo (1); Pino, Anna
CORPORATE SOURCE: (1) Lab. Comparative Toxicol. Ecotoxicol., Ist. Superiore Sanita, Viale Regina Elena 299-00161, Rome Italy
SOURCE: Mutation Research, (Oct. 12, 1998) Vol. 421, No. 1, pp. 93-107.
ISSN: 0027-5107.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:

The rodent carcinogenicity bioassay has been used for several decades for evaluating hundreds of chemicals, with the two aims of better understanding the etiologies of cancer, and of assessing the hazard posed by environmental and industrial chemicals. This has generated an enormous wealth of data and information on the phenomenon of chemical carcinogenicity. However, this information cannot be appreciated easily, since too many details may obscure the general trends present in the data; on the contrary, the use of computerized data analysis techniques suitable for the exploration of large databases makes its investigation much more fruitful, and its results more reliable. For this work, we collected a database of 536 rodent carcinogens, and we investigated the profiles of tumors (target organs) induced in the four experimental systems which are usually employed (rat and mouse, male and female). The analysis was performed with an Artificial Neural Network called Kohonen **Self-Organizing Map**, which is a computer-intensive method aimed at making the relevant information emerge automatically from the data itself. The analysis generated a global view, as well as a quantitative measure of the associations among the individual tumor types, and among the tumor profiles induced by the chemicals. In the complex interplay between the organ and species specificity of tumor induction, the species specificity generally overcame organ specificity, except for a few tumors (namely Lymphatic System, Brain, Forestomach, Stomach and Thyroid Gland). Moreover, the species specificity was remarkably stronger than the trans-species sex specificity. For three chemical classes (Aromatic Amines, Electrophilic/Alkylating Agents, Nitroarenes) most represented in the database, we investigated the hypothesis that a single mechanism of interaction with DNA would produce one, or a few very similar tumor profiles. Our analysis pointed out that no obvious association exists between chemical/mode of action class, and tumor profile. On the contrary, none of these classes induces a single tumor or **pattern** of tumors, but rather it appears that each class produces tumors at a wide range of sites. This suggests that an important determinant of the differences in tumor profile are the events that surround the ultimate mechanism of interaction with DNA.

CONCEPT CODE: Neoplasms and Neoplastic Agents - General *24002
Genetics and Cytogenetics - Animal *03506
Biochemical Studies - General *10060
Toxicology - General; Methods and Experimental *22501
BIOSYSTEMATIC CODE: Rodentia - Unspecified 86265
Muridae 86375
INDEX TERMS: Major Concepts
Toxicology; Tumor Biology
INDEX TERMS: Diseases
chemically-induced tumor: neoplastic disease, toxicity
INDEX TERMS: Chemicals & Biochemicals
aromatic amines; carcinogens; electrophilic/alkylating agents; nitroarenes; DNA
INDEX TERMS: Miscellaneous Descriptors
sex specificity; species specificity; structure-activity relationship; target organs; tumor induction; Kohonen **self-organizing map**: artificial neural network

ORGANISM:	Super Taxa Muridae: Rodentia, Mammalia, Vertebrata, Chordata, Animalia; Rodentia: Mammalia, Vertebrata, Chordata, Animalia
ORGANISM:	Organism Name mouse (Muridae): female, male; rat (Muridae): female, male; rodent (Rodentia): female, male
ORGANISM:	Organism Superterms Animals; Chordates; Mammals; Nonhuman Mammals; Nonhuman Vertebrates; Rodents; Vertebrates

L7 ANSWER 1 OF 5 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 2000:509680 BIOSIS
DOCUMENT NUMBER: PREV200000509680
TITLE: The McGill Pain Questionnaire in patients with TMJ pain and with facial pain as a somatoform disorder.
AUTHOR(S): Mongini, Franco (1); Italiano, Marco; Raviola, Fabio; Mossolov, Alexei
CORPORATE SOURCE: (1) Department of Clinical Pathophysiology, Unit of Headache and Facial Pain, University of Turin, Corso Dogliotti 14, I-10126, Torino Italy
SOURCE: Cranio, (October, 2000) Vol. 18, No. 4, pp. 249-256. print. ISSN: 0886-9634.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:
The purpose of this study was to assess the discriminative capacity of the McGill Pain Questionnaire (MPQ) in patients with temporomandibular joint disorders (TMD) or with facial pain disorder as somatoform disorder (referred to as "atypical facial pain") (FP). The MPQ was administered to 57 TMD and 34 FP patients. Weighted MPQ item scores, subscale Pain Rating Indexes (PRI), and total Pain Rating Index were tested for significant differences (Student's t-test), and the frequency of descriptor choice was also analyzed. Furthermore, the data were processed through two systems based on a counter-propagation neural network: the **Self-Organizing Map** (SOM) system and a **cluster**-like analysis. In the FP group eleven MPQ item scores and five PRI scores were significantly higher than those of the TMJ group. There was a considerable difference in descriptor choice between the groups. SOM analysis and **cluster**-like analysis correctly discriminated 85% or more of the patients. In conclusion, the MPQ showed a consistent discriminative capacity between TMD and FP patients.
CONCEPT CODE: Psychiatry - Psychopathology; Psychodynamics and Therapy *21002
Behavioral Biology - Human Behavior *07004
Nervous System - Pathology *20506
INDEX TERMS: Major Concepts
Neurology (Human Medicine, Medical Sciences); Methods and Techniques
INDEX TERMS: Diseases
somatoform disorder: behavioral and mental disorders
INDEX TERMS: Methods & Equipment
McGill Pain Questionnaire: evaluation method; pain rating indexes: evaluation method; total pain rating index: evaluation method
INDEX TERMS: Miscellaneous Descriptors
TMJ pain; facial pain
ORGANISM: Super Taxa
Hominidae: Primates, Mammalia, Vertebrata, Chordata, Animalia
ORGANISM: Organism Name
human (Hominidae): patient
ORGANISM: Organism Superterms
Animals; Chordates; Humans; Mammals; Primates; Vertebrates

L7 ANSWER 2 OF 5 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 2000:8353 BIOSIS
DOCUMENT NUMBER: PREV200000008353
TITLE: Neural network-based analysis of MR time series.
AUTHOR(S): Fischer, Harald (1); Hennig, Juergen
CORPORATE SOURCE: (1) Department of Radiology, University of Freiburg, Hugstetter Str. 55, D-79106, Freiburg Germany
SOURCE: Magnetic Resonance in Medicine, (Jan., 1999) Vol. 41, No. 1, pp. 124-131.
ISSN: 0740-3194.

DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:
Clustering has been introduced to analyze fMRI data by means of partitioning data into time series of similar temporal behavior. It is hoped that one of these clusters represents a dynamic effect of interest, like functional activation. Using **self-organizing** maps for clustering, additional information can be obtained by ordering **cluster** centers on a two-dimensional projection plane. The **map**'s capability of data visualization is used to summarize all dynamic effects of an experiment by means of data partitioning. The **map** does allow differently sized and populated clusters in the data by forming "superclusters" on the **map**. The method is introduced as a conceptual extension to clustering. Applications to fMRI and to MR mammography are discussed.
CONCEPT CODE: Radiation - General *06502
INDEX TERMS: Major Concepts
Methods and Techniques
INDEX TERMS: Methods & Equipment
MR mammography: imaging method; fMRI [functional magnetic resonance imaging]: imaging method; neural network-based analysis: analytical method; **self-organizing map**: imaging method
INDEX TERMS: Miscellaneous Descriptors
data visualization

L7 ANSWER 3 OF 5 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1999:277453 BIOSIS
DOCUMENT NUMBER: PREV199900277453
TITLE: Analysis of gene expression data using **self-organizing** maps.
AUTHOR(S): Toronen, Petri; Kolehmainen, Mikko; Wong, Garry; Castren, Eero (1)
CORPORATE SOURCE: (1) A.I. Virtanen Institute, University of Kuopio, 70211, Kuopio Finland
SOURCE: FEBS Letters, (May 21, 1999) Vol. 451, No. 2, pp. 142-146. ISSN: 0014-5793.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:
DNA microarray technologies together with rapidly increasing genomic sequence information is leading to an explosion in available gene expression data. Currently there is a great need for efficient methods to analyze and visualize these massive data sets. A **self-organizing map** (SOM) is an unsupervised neural network learning algorithm which has been successfully used for the analysis and organization of large data files. We have here applied the SOM algorithm to analyze published data of yeast gene expression and show that SOM is an excellent tool for the analysis and visualization of gene expression profiles.
CONCEPT CODE: Genetics and Cytogenetics - Plant *03504
Mathematical Biology and Statistical Methods *04500
Replication, Transcription, Translation *10300
Plant Physiology, Biochemistry and Biophysics - Metabolism *51519
Plant Physiology, Biochemistry and Biophysics - Apparatus and Methods *51524
Plant Physiology, Biochemistry and Biophysics - General and Miscellaneous *51526
BIOSYSTEMATIC CODE: Fungi - Unspecified 15000
INDEX TERMS: Major Concepts
Genetics; Mathematical Biology (Computational Biology); Methods and Techniques
INDEX TERMS: Methods & Equipment

cluster analysis: mathematical method;
self-organizing map: analytical
method, mathematical method

INDEX TERMS: Miscellaneous Descriptors
gene expression analysis

ORGANISM: Super Taxa
Fungi: Plantae

ORGANISM: Organism Name
yeast (Fungi)

ORGANISM: Organism Superterms
Fungi; Microorganisms; Nonvascular Plants; Plants

L7 ANSWER 4 OF 5 BIOSIS COPYRIGHT 2001 BIOSIS

ACCESSION NUMBER: 1998:91913 BIOSIS

DOCUMENT NUMBER: PREV199800091913

TITLE: Feature-extraction from endopeptidase cleavage sites in mitochondrial targeting peptides.

AUTHOR(S): Schneider, Gisbert; Sjoling, Sara; Wallin, Erik; Wrede, Paul; Glaser, Elzbieta; Von Heijne, Gunnar (1)

CORPORATE SOURCE: (1) Dep. Biochem., Stockholm Univ., S-10691 Stockholm Sweden

SOURCE: Proteins Structure Function and Genetics, (Jan. 1, 1998) Vol. 30, No. 1, pp. 49-60.
ISSN: 0887-3585.

DOCUMENT TYPE: Article

LANGUAGE: English

ABSTRACT:

Cleavage sites in nuclear-encoded mitochondrial protein targeting peptides (mTPs) from mammals, yeast, and plants have been analysed for characteristic physicochemical features using statistical methods, perceptrons, multilayer neural networks, and **self-organizing** feature maps. Three different sequence motifs were found, revealing loosely defined arginine motifs with Arg in positions -10, -3, and -2. A **self-organizing** feature **map** was able to **cluster** these three types of endopeptidase target sites but did not identify any species-specific characteristics in mTPs. Neural networks were used to define local sequence features around precursor cleavage sites.

CONCEPT CODE: Enzymes - General and Comparative Studies; Coenzymes *10802
Biochemical Studies - General *10060
Biophysics - General Biophysical Studies *10502

BIOSYSTEMATIC CODE: Fungi - Unspecified 15000
Ascomycetes 15100
Gramineae 25305
Leguminosae 26260
Bovidae 85715
Suidae 85740
Hominidae 86215
Muridae 86375

INDEX TERMS: Major Concepts
Enzymology (Biochemistry and Molecular Biophysics)

INDEX TERMS: Parts, Structures, & Systems of Organisms
mitochondria

INDEX TERMS: Chemicals & Biochemicals
mitochondrial intermediate peptidase; mitochondrial processing peptidase; mitochondrial protein targeting peptide: endopeptidase cleavage sites, molecular structure

INDEX TERMS: Methods & Equipment
multilayer neural networks: analytical method; perceptrons: analytical method; **self-organizing**
feature maps: analytical method

INDEX TERMS: Miscellaneous Descriptors
statistical methods

ORGANISM: Super Taxa

Ascomycetes: Fungi, Plantae; Bovidae: Artiodactyla, Mammalia, Vertebrata, Chordata, Animalia; Fungi: Plantae; Gramineae: Monocotyledones, Angiospermae, Spermatophyta, Plantae; Hominidae: Primates, Mammalia, Vertebrata, Chordata, Animalia; Leguminosae: Dicotyledones, Angiospermae, Spermatophyta, Plantae; Muridae: Rodentia, Mammalia, Vertebrata, Chordata, Animalia; Suidae: Artiodactyla, Mammalia, Vertebrata, Chordata, Animalia

ORGANISM: Organism Name
cow (Bovidae); human (Hominidae); maize (Gramineae); mouse (Muridae); pea (Leguminosae); pig (Suidae); rat (Muridae); yeast (Fungi); Neurospora-crassa [yeast] (Ascomycetes)

ORGANISM: Organism Superterms
Angiosperms; Animals; Artiodactyls; Chordates; Dicots; Fungi; Humans; Mammals; Microorganisms; Monocots; Nonhuman Mammals; Nonhuman Vertebrates; Nonvascular Plants; Plants; Primates; Rodents; Spermatophytes; Vascular Plants; Vertebrates

REGISTRY NUMBER: 9001-92-7 (ENDOPEPTIDASE)
9031-96-3 (PEPTIDASE)

L7 ANSWER 5 OF 5 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1994:104157 BIOSIS
DOCUMENT NUMBER: PREV199497117157
TITLE: Potentially functional regions of nucleic acids recognized by a Kohonen's **self-organizing map.**
AUTHOR(S): Giuliano, F.; Arrigo, P.; Scalia, F.; Cardo, P. P.; Damiani, G. (1)
CORPORATE SOURCE: (1) Istituto Policattedra di Chimica Biologica, Viale Benedetto XV 1, 16232 Genova Italy
SOURCE: Computer Applications in the Biosciences, (1993) Vol. 9, No. 6, pp. 687-693.
ISSN: 0266-7061.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:
Computer recognition of short functional sites on DNA, such as promoter regions or intron-exon boundaries, has recently attracted much interest. In this paper we have focused our attention on the automatic recognition of relevant features of human nucleic acid sequences by means of an unsupervised artificial neural network model. Sixty messenger RNA and 31 genomic DNA sequences were analysed. The results showed that in mRNA, the minimal similarity 60 base pattern was guanine- and cytosine-rich and located in most sequences in a range of 250 bases from either the middle point of the signal peptide coding region or from the start of the coding region. On DNA sequences a region defined by a ***cluster*** of minimal similarity patterns was present in many of the analysed genes. This zone may be related to alternative splicing and DNA methylation.
CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Nucleic Acids, Purines and Pyrimidines *10052
Biochemical Studies - Nucleic Acids, Purines and Pyrimidines 10062
Biophysics - Molecular Properties and Macromolecules *10506
Nervous System - General; Methods *20501

INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Information Studies; Mathematical Biology (Computational Biology); Methods and Techniques; Nervous System (Neural Coordination)

INDEX TERMS: Sequence Data

nucleotide sequence

INDEX TERMS: Miscellaneous Descriptors
ARTIFICIAL NEURAL NETWORK; COMPUTER ANALYSIS; DNA; EMBL; METHYLATION; SPLICING

L1 ANSWER 1 OF 1 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1991:431331 BIOSIS
DOCUMENT NUMBER: BA92:87496
TITLE: IDENTIFICATION OF A NEW MOTIF ON **NUCLEIC** ACID SEQUENCE DATA USING KOHONEN'S **SELF-ORGANIZING MAP**.
AUTHOR(S): ARRIGO P; GIULIANO F; SCALIA F; RAPALLO A; DAMIANI G
CORPORATE SOURCE: IST. I CIRCUITI ELETTRONICI C.N.R., VIA ALL'OPERA PIA 11, 16145 GENOVA, ITALY.
SOURCE: COMPUT APPL BIOSCI, (1991) 7 (3), 353-358.
CODEN: COABER. ISSN: 0266-7061.
FILE SEGMENT: BA; OLD
LANGUAGE: English
ABSTRACT:

Here we present a performance test of a Kohonen features **map** applied to the fast extraction of uncommon sequences from the coding region of the human insulin receptor **gene**. We used a network with 30 neurons and with a variable input window. The program was aimed at detecting unique or uncommon DNA regions present in crude sequence data and was able to automatically detect the signal peptide coding regions of a set of human insulin receptor **gene** data. The testing of this programn with HSIRPR cDNA release (EMBL data bank) indicated the presence of unique features in the signal peptide coding region. On the basis of our results this program can automatically detect 'singularity' from crude sequencing data and it does not require knowledge of the features to be found.

CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Methods, Materials and Apparatus, General - Laboratory Apparatus 01006
Genetics and Cytogenetics - Human *03508
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Nucleic Acids, Purines and Pyrimidines 10052
Biophysics - Bioengineering 10511
Biophysics - Biocybernetics *10515
Metabolism - Nucleic Acids, Purines and Pyrimidines *13014
Endocrine System - Pancreas *17008
Psychiatry - Mental Retardation 21006
BIOSYSTEMATIC CODE: Hominidae 86215
INDEX TERMS: Miscellaneous Descriptors
HUMAN INSULIN RECEPTOR **GENE** DATA
REGISTRY NUMBER: 9004-10-8 (INSULIN)

L2 ANSWER 1 OF 2 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1999:277453 BIOSIS
DOCUMENT NUMBER: PREV199900277453
TITLE: Analysis of **gene** expression data using **self-organizing** maps.
AUTHOR(S): Toronen, Petri; Kolehmainen, Mikko; Wong, Garry; Castren, Eero (1)
CORPORATE SOURCE: (1) A.I. Virtanen Institute, University of Kuopio, 70211, Kuopio Finland
SOURCE: FEBS Letters, (May 21, 1999) Vol. 451, No. 2, pp. 142-146. ISSN: 0014-5793.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:
DNA microarray technologies together with rapidly increasing genomic sequence information is leading to an explosion in available **gene** expression data. Currently there is a great need for efficient methods to analyze and visualize these massive data sets. A **self-organizing** ***map*** (SOM) is an unsupervised neural network learning algorithm which has been successfully used for the analysis and organization of large data files. We have here applied the SOM algorithm to analyze published data of yeast **gene** expression and show that SOM is an excellent tool for the analysis and visualization of **gene** expression profiles.
CONCEPT CODE: Genetics and Cytogenetics - Plant *03504
Mathematical Biology and Statistical Methods *04500
Replication, Transcription, Translation *10300
Plant Physiology, Biochemistry and Biophysics - Metabolism *51519
Plant Physiology, Biochemistry and Biophysics - Apparatus and Methods *51524
Plant Physiology, Biochemistry and Biophysics - General and Miscellaneous *51526
BIOSYSTEMATIC CODE: Fungi - Unspecified 15000
INDEX TERMS: Major Concepts
Genetics; Mathematical Biology (Computational Biology); Methods and Techniques
INDEX TERMS: Methods & Equipment
cluster analysis: mathematical method; **self-organizing map**: analytical method, mathematical method
INDEX TERMS: Miscellaneous Descriptors
gene expression analysis
ORGANISM: Super Taxa
Fungi: Plantae
ORGANISM: Organism Name
yeast (Fungi)
ORGANISM: Organism Superterms
Fungi; Microorganisms; Nonvascular Plants; Plants

L2 ANSWER 2 OF 2 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1991:431331 BIOSIS
DOCUMENT NUMBER: BA92:87496
TITLE: IDENTIFICATION OF A NEW MOTIF ON NUCLEIC ACID SEQUENCE DATA USING KOHONEN'S **SELF-ORGANIZING MAP**.
AUTHOR(S): ARRIGO P; GIULIANO F; SCALIA F; RAPALLO A; DAMIANI G
CORPORATE SOURCE: IST. I CIRCUITI ELETTRONICI C.N.R., VIA ALL'OPERA PIA 11, 16145 GENOVA, ITALY.
SOURCE: COMPUT APPL BIOSCI, (1991) 7 (3), 353-358.
CODEN: COABER. ISSN: 0266-7061.
FILE SEGMENT: BA; OLD
LANGUAGE: English
ABSTRACT:

Here we present a performance test of a Kohonen features **map** applied to the fast extraction of uncommon sequences from the coding region of the human insulin receptor **gene**. We used a network with 30 neurons and with a variable input window. The program was aimed at detecting unique or uncommon DNA regions present in crude sequence data and was able to automatically detect the signal peptide coding regions of a set of human insulin receptor **gene** data. The testing of this programn with HSIRPR cDNA release (EMBL data bank) indicated the presence of unique features in the signal peptide coding region. On the basis of our results this program can automatically detect 'singularity' from crude sequencing data and it does not require knowledge of the features to be found.

CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Methods, Materials and Apparatus, General - Laboratory Apparatus 01006
Genetics and Cytogenetics - Human *03508
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Nucleic Acids, Purines and Pyrimidines 10052
Biophysics - Bioengineering 10511
Biophysics - Biocybernetics *10515
Metabolism - Nucleic Acids, Purines and Pyrimidines *13014
Endocrine System - Pancreas *17008
Psychiatry - Mental Retardation 21006
BIOSYSTEMATIC CODE: Hominidae 86215
INDEX TERMS: Miscellaneous Descriptors
HUMAN INSULIN RECEPTOR **GENE** DATA
REGISTRY NUMBER: 9004-10-8 (INSULIN)

L4 ANSWER 1 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 2000:242442 BIOSIS
DOCUMENT NUMBER: PREV200000242442
TITLE: The effect of intracortical competition on the formation of topographic maps in models of Hebbian learning.
AUTHOR(S): Piepenbrock, C.; Obermayer, K. (1)
CORPORATE SOURCE: (1) Fachbereich Informatik, Technische Universitaet Berlin, Franklinstrasse 28/29, FR2-1, D-10587, Berlin Germany
SOURCE: Biological Cybernetics, (April, 2000) Vol. 82, No. 4, pp. 345-353.
ISSN: 0340-1200.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English
ABSTRACT:

Correlation-based learning (CBL) models and **self-organizing** maps (SOM) are two classes of Hebbian models that have both been proposed to explain the activity-driven formation of cortical maps. Both models differ significantly in the way lateral cortical interactions are treated, leading to different predictions for the formation of receptive fields. The linear CBL models predict that receptive field profiles are determined by the average values and the spatial correlations of the second order of the afferent activity patterns, whereas SOM models **map** stimulus features. Here, we investigate a class of models which are characterized by a variable degree of lateral competition and which have the CBL and SOM models as limit cases. We show that there exists a critical value for intracortical competition below which the model exhibits CBL properties and above which feature mapping sets in. The class of models is then analyzed with respect to the formation of topographic maps between two layers of neurons. For Gaussian input stimuli we find that localized receptive fields and topographic maps emerge above the critical value for intracortical competition, and we calculate this value as a function of the size of the input stimuli and the range of the lateral interaction function. Additionally, we show that the learning rule can be derived via the optimization of a global cost function in a framework of probabilistic output neurons which represent a set of input stimuli by a sparse code.

CONCEPT CODE: Nervous System - General; Methods *20501
Mathematical Biology and Statistical Methods *04500
Biophysics - Biocybernetics *10515
INDEX TERMS: Major Concepts
Models and Simulations (Computational Biology); Nervous System (Neural Coordination)
INDEX TERMS: Parts, Structures, & Systems of Organisms
neurons: nervous system
INDEX TERMS: Miscellaneous Descriptors
Hebbian learning models: applications; **biological** cybernetics; correlation-based learning models: applications; intracortical competition; mathematical models: applications; **self-organizing** maps; topographical maps: formation

L4 ANSWER 3 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1999:45081 BIOSIS
DOCUMENT NUMBER: PREV199900045081
TITLE: Comparison of chemical databases: Analysis of molecular diversity with **self organizing** maps (SOM.
AUTHOR(S): Bernard, P. (1); Golbraikh, A. (1); Kireev, D. (1); Chretien, J. R. (1); Rozhkova, N.
CORPORATE SOURCE: (1) Lab. Chemometrics, Univ. Orleans, BP 6759, 45067 Orleans Cedex 2 France
SOURCE: Analusis, (Oct., 1998) Vol. 26, No. 8, pp. 333-341. ISSN: 0365-4877.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:

Self Organising **Map** (SOM), also known as Kohonen Neural Network, is tested as a non supervised procedure for comparing molecular databases. Each chemical compound being represented by a point in the hyperspace of the molecular descriptors, SOMs was used to reflect the multidimensional hyperspace onto a two dimensional (2D) **map** while preserving the order of distances between the points, but in a non linear way. The aim of this work was to apply SOM to the study of the overlapping of two databases in order to obtain information about the extent of their differences in regard to their molecular diversity. Firstly, the ability of SOM to discriminate between two virtual databases was investigated. The positions of these two virtual databases were made to vary from non-overlapping to overlapping ones. In any considered cases, all the individuals of these two databases are processed simultaneously to give one SOM. From this **map** it is possible to analyse and understand the structure of the original data. Secondly two chemical databases are compared. The first chemical database deals with the commercially available organophosphorous pesticides (OPC), the second one deals with more than two thousand OPC tested as potent pesticides. Given the ***biological*** data known for each compound, the second database was shown to bring an interesting supplement to the structural information nested in the first database taken as a reference. Furthermore, the results obtained indicate that SOM can be used for the search of new leads among available databases and the exploration of new structural domains for a given **biological** activity.

CONCEPT CODE: Pest Control, General; Pesticides; Herbicides *54600
Mathematical Biology and Statistical Methods *04500
Biochemical Studies - General *10060
INDEX TERMS: Major Concepts
Information Studies; Methods and Techniques
INDEX TERMS: Chemicals & Biochemicals
organophosphorous pesticides
INDEX TERMS: Methods & Equipment
self organizing maps:
Analysis/Characterization Techniques: CB, analytical method
INDEX TERMS: Miscellaneous Descriptors
biological activity; chemical databases; molecular diversity

L4 ANSWER 4 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1997:502350 BIOSIS
DOCUMENT NUMBER: PREV199799801553
TITLE: Classification of protein families and detection of the determinant residues with an improved **self-organizing map.**
AUTHOR(S): Andrade, Miguel A. (1); Casari, Georg; Sander, Chris; Valencia, Alfonso
CORPORATE SOURCE: (1) European Bioinformatics Inst,. Hinxton, Cambridge CBIO 1SD UK
SOURCE: Biological Cybernetics, (1997) Vol. 76, No. 6, pp. 441-450. ISSN: 0340-1200.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:

Using a SOM (**self-organizing map**) we can classify sequences within a protein family into subgroups that generally correspond to ***biological*** subcategories. These maps tend to show sequence similarity as proximity in the **map**. Combining maps generated at different levels of resolution, the structure of relations in protein families can be captured that could not otherwise be represented in a single **map**. The underlying representation of maps enables us to retrieve characteristic sequence patterns for individual subgroups of sequences. Such patterns tend to correspond to functionally important regions. We present a modified SOM algorithm that includes a convergence test that dynamically controls the learning parameters to adapt them to the learning set instead of being fixed and externally optimized by trial and error. Given the variability of protein family size and distribution, the addition of this feature is necessary. The method is successfully tested with a number of families. The rab family of small GTPases is used to illustrate the performance of the method.

CONCEPT CODE: Mathematical Biology and Statistical Methods *04500
Biochemical Methods - Proteins, Peptides and Amino Acids *10054
Biochemical Studies - Proteins, Peptides and Amino Acids *10064
Biophysics - Molecular Properties and Macromolecules *10506
Biophysics - Biocybernetics *10515
INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Mathematical Biology (Computational Biology); Methods and Techniques; Models and Simulations (Computational Biology)
INDEX TERMS: Miscellaneous Descriptors
BIOCYBERNETICS; CLASSIFICATION; CLUSTERING ALGORITHM; DETERMINANT RESIDUE DETECTION; MODELS AND SIMULATIONS; PROTEIN FAMILY; **SELF-ORGANIZING MAP**

L4 ANSWER 5 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1997:385433 BIOSIS
DOCUMENT NUMBER: PREV199799684636
TITLE: Coordinate-free self-organising feature maps.
AUTHOR(S): Zuzan, Harry; Holbrook, John A.; Kim, Peter T.; Harauz, George (1)
CORPORATE SOURCE: (1) Dep. Mol. Biol. Genetics, Univ. Guelph, Guelph, ON N1G 2W1 Canada
SOURCE: Ultramicroscopy, (1997) Vol. 68, No. 3, pp. 201-214. ISSN: 0304-3991.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:

The successful application of a new strategy for classifying images of ***biological*** macromolecules, and resolving their rotational orientations, was recently introduced by R. Marabini and J. M. Carazo (Pattern recognition and classification of images of **biological** macromolecules using artificial neural networks, Biophys. J. 66 (1994) 1801-1814). Their work was based on Kohonen's **self-organizing** features **map** (SOFM) defined on a plane, and has been extended here by allowing an SOFM to operate independently of topology. An SOFM has been constructed which follows instructions according to the current values of a variable, which alone drive the **self-organizing** process. The instructions that the SOFM follows and only available internally to the **map** and so the behaviour of the SOFM must be supervised by providing suggestions as to what the state of its components should be. The method is shown to be useful in identification and clustering of recurring motifs, of resolving metastable states in which the process can occasionally become trapped, and in discarding data unsuitable for further analysis.

CONCEPT CODE: Microscopy Techniques - General and Special Techniques *01052
Biochemical Methods - General *10050
Biophysics - Molecular Properties and Macromolecules *10506
INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Methods and Techniques
INDEX TERMS: Miscellaneous Descriptors
ANALYTICAL METHOD; BIOCHEMISTRY AND BIOPHYSICS; **BIOLOGICAL** MACROMOLECULES; COORDINATE-FREE **SELF-ORGANIZING** FEATURE MAPS; MACROMOLECULAR MICROSCOPY; METHODOLOGY

L4 ANSWER 7 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1997:21025 BIOSIS
DOCUMENT NUMBER: PREV199799320228
TITLE: Analysis of structural variability within two-dimensional **biological** crystals by a combination of patch averaging techniques and **self organizing** maps.
AUTHOR(S): Fernandez, Jose-Jesus; Carazo, Jose-Maria (1)
CORPORATE SOURCE: (1) Cent. Nacl. Biotecnol. CSIC, Univ. Autonoma, 28049 Madrid Spain
SOURCE: Ultramicroscopy, (1996) Vol. 65, No. 1-2, pp. 81-93. ISSN: 0304-3991.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:
We study in this work the use of **self organizing** maps to analyze the structural variability that can be found along two-dimensional crystals of **biological** macromolecules. Small areas of the crystals, termed "patches" by previous researchers, are used to obtain local average images that are then used as the input of a **Self Organizing** ***Map*** . This procedure allows for a fast and accurate image classification. Multivariate Statistical Analysis is then used on the resulting code vectors producing a very condensed data representation. This methodology is applied to previously studied crystals of bacteriophage vphi-29 p10 connector, finding a crystalline heterogeneity probably associated to multilayers in some areas of the crystal.
CONCEPT CODE: General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Mathematical Biology and Statistical Methods *04500
Biochemical Methods - General *10050
Biochemical Methods - Proteins, Peptides and Amino Acids *10054
Biochemical Studies - General *10060
Biochemical Studies - Proteins, Peptides and Amino Acids *10064
Virology - Bacteriophage *33504
INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Information Studies; Mathematical Biology (Computational Biology); Methods and Techniques; Microbiology
INDEX TERMS: Miscellaneous Descriptors
BACTERIOPHAGE-PHI-29 P10 CONNECTOR; BIOCHEMISTRY AND BIOPHYSICS; **BIOLOGICAL** MACROMOLECULE; COMPUTER APPLICATIONS; COMPUTER PROGRAMS; MATHEMATICAL BIOLOGY; MATHEMATICAL METHOD; MATHEMATICAL METHODS; MULTIVARIATE STATISTICAL ANALYSIS; NEURAL NETWORK; PATCH AVERAGING TECHNIQUES; **SELF ORGANIZING MAP** ALGORITHM; STRUCTURAL VARIABILITY ANALYSIS; TWO-DIMENSIONAL CRYSTAL

L4 ANSWER 8 OF 8 BIOSIS COPYRIGHT 2001 BIOSIS
ACCESSION NUMBER: 1994:310166 BIOSIS
DOCUMENT NUMBER: PREV199497323166
TITLE: Pattern recognition and classification of images of **biological** macromolecules using artificial neural networks.
AUTHOR(S): Marabini, R.; Carazo, J. M. (1)
CORPORATE SOURCE: (1) Centro Nacional Biotechnologia, Univ. Autonoma, Campus de Canto Blanco, 28049 Madrid Spain
SOURCE: Biophysical Journal, (1994) Vol. 66, No. 6, pp. 1804-1814. ISSN: 0006-3495.
DOCUMENT TYPE: Article
LANGUAGE: English
ABSTRACT:

The goal of this work was to analyze an image data set and to detect the structural variability within this set. Two algorithms for pattern recognition based on neural networks are presented, one that performs an unsupervised classification (the **self-organizing map**) and the other a supervised classification (the learning vector quantization). The approach has a direct impact in current strategies for structural determination from electron microscopic images of **biological** macromolecules. In this work we performed a classification of both aligned but heterogeneous image data sets as well as basically homogeneous but otherwise rotationally misaligned image populations, in the latter case completely avoiding the typical reference dependency of correlation-based alignment methods. A number of examples on chaperonins are presented. The approach is computationally fast and robust with respect to noise. Programs are available through ftp.

CONCEPT CODE:
General Biology - Information, Documentation, Retrieval and Computer Applications *00530
Microscopy Techniques - Electron Microscopy *01058
Mathematical Biology and Statistical Methods *04500
Comparative Biochemistry, General *10010
Biochemical Methods - General *10050
Biochemical Methods - Proteins, Peptides and Amino Acids *10054
Biochemical Studies - General *10060
Biochemical Studies - Proteins, Peptides and Amino Acids *10064
Biophysics - General Biophysical Studies *10502
Biophysics - Molecular Properties and Macromolecules *10506
Biophysics - Biocybernetics *10515

INDEX TERMS: Major Concepts
Biochemistry and Molecular Biophysics; Information Studies; Mathematical Biology (Computational Biology); Methods and Techniques; Models and Simulations (Computational Biology)

INDEX TERMS: Miscellaneous Descriptors
CHAPERONINS; COMPUTATION SPEED; COMPUTER PROGRAMS; ELECTRON MICROSCOPY; MOLECULAR STRUCTURE; STRUCTURAL VARIABILITY